E-ISSN 2281-4612 | Academic Journal of Interdisciplinary Studies | Vol 2 No 10
ISSN 2281-3993 | MCSER Publishing, Rome-Italy | October 2013

Cultural Imperatives in Differential Item Functioning (DIF) in Mathematics

Ebisine, Sylvester Sele

*Department of Mathematics, College of Education,
Warri, Delta State,Nigeria*

Doi:10.5901/ajis.2013.v2n10p85

Abstract

DIF is said to occur if different subgroups, who have equal standing on the construct the test is designed to measure, display different probabilities of passing an item (in test of typical performance). DIF as a condition in which given examinees of equal ability in the trait being measured, the probability of answering an item correctly is related to group membership. There are evidences of DIF among testees from different subgroups (cultural group, gender and socioeconomic background). Therefore, any test used to measure students' ability must be fair among different cultures. Because of the different cultural setting tests may not be functioning in the same way in all cultures, which is also called that test may be equivalent or tests may not be fair among different culture. This paper therefore discusses the concepts of DIF, culture and the theoretical perspective of cultural imperative in DIF. Furthermore, the paper examined the sources of DIF in multicultural assessment and how to deal with inequivalent items. Finally, it attempts the use of culturally responsive instructional technique to reduce students' differential performance in mathematics.

Keywords: *Culture, Mathematics, Differential Item Functioning (DIF), Inequivalent Items, Differential Performance*

1. Introduction

Learning is a basic attribute of every human being. No learning is complete until what is learnt is assessed. The process of determining the extent of possession of what was learnt in a concrete and measurable term is called assessment. A basic component of assessment is measurement. Measurement is the transformation of attributes to numeric values. Gronlund (1976) describe measurement as the quantitative description of learner's behavior. In psychological measures, because the attribute being measured is not directly observable it is given a special name "construct". Then, the equivalence of the attribute measured can be re-specified as: the construct to be measured must possess the same properties and meaning in different culture, which is referred to as construct equivalence (Hui & Triandis, 1983). In addition, the scale representing this construct must also be the same. Such an identical scale is referred to as comparison scale. In other words, comparison scale can be regarded as a measurement scale on which equivalence is assured.

However, the problem due to the lack of equivalence in the construct being measured is called construct bias or item bias. Item bias may be due to various reasons, such as partial correspondence of the construct over culture. If the construct being measured does not possess the same set of behaviours in different culture, presence of irrelevant, non-target constructs which are related to gender, ethnicity, race, linguistic background, socio-economic status or handicapping condition (Flores, 2000; Lam, 1995), differences in upbringing environment, culture (Flores, 2000) and daily life experiences (Fortune, 1985). The general concept is termed Differential Item

Functioning (DIF). DIF is said to occur if different subgroups, who have equal standing on the construct the test is designed to measure, display different probabilities of passing an item (in test of typical performance).

Therefore, any test used to measure students' ability must be fair among different cultures (Poortinga, 1989; Kleime & Baumert, 2001). In other words, it must be assured that the tests measure some common construct in different cultures. Because of the different cultural setting tests may not be functioning in the same way in all cultures, which is also called that test may be equivalent or tests may not be fair among different culture (Allalouff, Hambleton & Sireci, 1999; Ercikan, 1998). For example, cultural differences may lead to different response styles or response patterns in examination. Presence of these factors with the potential of affecting item equivalence can cause problems in comparability of items (Sireci & Berberoglu, 2000; Arim & Ercikan, 2005). For example, the Ijaw cultural environment favour spatial component of mathematics as reflected in geometry, whereas the Hausa cultural environment allied with the numerical component of mathematics which is arithmetic in nature. For instance, in NECO (2011) multiple choice test on mathematics, 33% of the items were phrased in cultural environment. Questions 3, 4, 11, 15, 53, 54, 58 favour the Hausa culture which is trading in nature. While questions 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 48 and 49 favour the spatial component of mathematics as reflected in the Ijaw cultural environment where activities are spatial in nature. As such test items in this aspect of mathematics may be simpler to an individual from such cultural environment. Specifically, question 3 which states as follows: find the compound interest on ₦225.00 invested for 20 years at 4% per annum favour the Hausa trade culture, while question 34 which states as follows: a right circular cylinder of base radius 8cm has a height of 14cm. find the volume of the cylinder favour the Ijaw spatial culture. Thus, if a test aims at measuring ability across different cultures, it has to assure that the scale on which the score are expressed has the same zero-point and the same scale units. This is to assume completely bias-free measurement. Item bias can seriously threaten the scalar equivalence.

Based on this background, this paper attempt to examine the concept of DIF, culture, sources of DIF in multi-cultural assessment and how to deal with inequivalent items. Finally, the paper examine how culturally responsive instructional technique can reduce students differential performance in mathematics.

2. Concept of Differential Item Functioning (DIF)

DIF occurs when a test item measure an ability which is alien to the subject matter, such that performance on the item is now sustained by abilities which are alien to the subject matter. According to Zumbo (1999) DIF is a condition in which examinees from different groups show differential probabilities of endorsing an item after matching on the underlying ability that the item is intended to measure. Put in a similar way Doolittle & Cleary (1987) defined DIF as a condition in which given examinees of equal ability in the trait being measured, the probability of answering an item correctly is related to group membership. In probability term, Warm (1978) defined DIF with the equation below:

$$PA\ (\Theta = K) \neq PB\ (\Theta = K)$$

In the equation above, A and B represent two subgroups, Θ represents ability which is equal to K in each group membership. The equation means that an item is said to be differentially functioning if probability of members of group A ability (K) getting an item correct is not equal to the probabilities of group B with ability K also getting the item correct. DIF comes about when different subgroups that are matched in terms of the primary dimension differ in their standing on a secondary dimension, such as familiarity with the content in terms of which the items are formulated. DIF analysis are directed at identifying items that are affected by secondary dimensions or, stated differently, that measure different, additional aspects in different subgroups. Nowadays, this term is invoked only if items have been identified, by statistical means, as differentially functioning, and if the reason for this can be attributed to construct-irrelevant properties of the item (Lam, 1995).

Similarly, Differential Person Function (DPF) indicates the differential person functioning between items and classes of persons. Another type of DIF is Differential Group Functioning (DGF), which indicates the differential functioning between classes of items and classes of persons, DIF and DPF.

3. What is Culture?

Culture comprises the distinctive habit of a people; it performs both a unifying and, more importantly a directive role. Culture referred specifically to those habits which bind a people together into a single group for a common end. According to Guillerman & Sharon (2001) culture and society shape an individual's mind and thinking. In effect the socio-cultural context in which people live influences the way in which they make sense of issues and the way in which they solve them. These socio-cultural influences include the values, beliefs, experiences, communication patterns, teaching and learning styles, and epistemologies inherent in the individual's cultural backgroups, as well as, the socio-economic conditions prevailing in cultural groups.

Reflecting on a cognitive orientation, Greenfield (1997) defined culture as the collective programming of the mind which distinguishes the members of one human group from another. Also focusing on cognition, Flores (2000) referred to the human constructed intangible in the environment as the subjective culture, defining it as a shared meaning system, founding among those who speak a particular language dialect, during a specific historic period and in definable geographic region. In essence, culture determines what individuals sense and perceive within their environment. It determines the importance of competing external and internal stimuli, the interpretation of the stimuli, and the consequential responses to the interpretations. Therefore, culture is an essential variable to consider when attempting to explain, alter, or predict behavior, as well as, performance difference between two groups of comparable ability or performance.

4. Theoretical Framework

This paper focus is on the interactionist view. The interactionist view focus is not the individual but interaction between individuals within a culture (Brunner, 1984). In essence for an interactionist mathematics educator, learning is not just an endeavor of the individual mind trying to adapt to an environment, nor can it be reduced to a process of enculturation into a pre-established culture. In mathematics, the individual construction of meanings takes place in interaction with the culture of environment (Cobb & Bauersfeld, 1995). Therefore, people learn indirectly, through participating in a culture and its discursive practices. For example, the process of construction of knowledge is based on interpretations that have their sources not in the individual alone but in his/her interpretation with others within a culture. For Bauersfeld (1995) meanings are generated neither by the individual mind nor are they attributed to some historically founded collective mind of a society, but they are continually constituted in the interactions whose pattered character accounts for the relative stability of cultures. Thus, the quality of the culture in which one lives play a dominant role in performance and development of mathematics knowledge.

The perspective focused on the fact that there are mathematical thinking used by students to solve problems that arise in customary activity such as their play, trade, farm activities and work (Schliemann, Carraher & Ceci, 1997). For the interactionist, mathematical thinking used by children to solve problems, could by the intellectual tools children acquire from their culture, and other support parents and teachers provide that help children develop mathematical knowledge (Radziszewska & Rogoff, 1991). This perspective indicates that variables associated with culture play a role in mathematical development. Although, it remains an open question whether or how culture per se play a role in mathematical development.

5. Source of DIF in Multi-Cultural Assessment

For almost past quarter of a century it has been recognized that DIF is caused by multi-dimensionality in an item (Linn, Levine, Harstings & Wardrop, 1980), that is, the performance on items depends not only on the construct that the test is designed to measure, referred to as primary (or target) dimension, but also on one or more dimensions, known as secondary dimension. Allalouf, Hambleton & Sireci (1999) provided one of the most comprehensive classification of causes of DIF in translated verbal items. They reported four main causes for DIF in translated instruments: (1) changes in difficulty of words or sentences specifies the situation in which some words become easier or more difficult after translation, (2) changes in contest may be due to an incorrect translation changing the meaning of an item. Gierl & Khalig's (2000) category of "omission or addition that affect meaning" also deals with same issues, (3) changes in format are the cases, for example when a sentence become much longer after the translation. Gierl & Khalig (2000) also include the changes in punctuation, capitalization in this category, (4) differences in cultural relevance is the last category. In this case the items remain same, however it is the cultural content of items that causes DIF. For example, content of a sentence completion item may be more familiar for one of the groups. Gierl & Khalig define this category as "differences in words or expressions inherent to a language". Scheuneman & Grima (1997), on the other hand provides an additional cognitive perspective to the classification of possible sources of group performance differences in mathematics items. They specify three categories, mainly (1) the cognitive nature of the task presented to the examinee, (2) mathematical content of the item, and (3) the surface properties of the items such as item format among others.

6. How to Deal with Inequivalent Items

Biased items are eliminated through the reduction of equivalence. For example, Hulin (1987) explained a method of eliminating the non-equivalent items and then reestimating their abilities with the rest of the items and again testing all items for equivalence and eliminating non-equivalent items and so on until no non-equivalent items are found. On the other hand, Roznowski & Reith (1999) provide an interesting perspective with result of their study that is not supporting the assumption of differentially functioning items should be deleted in order to get a fair measurement. They suggest that after determining the differentially functioning item it must be the following concern to test whether elimination of these items contributed to the quality of the test, because differential item functioning is an item level analysis whereas individuals are usually compared at the test level. Also, biased items may be pointing out some cross-cultural differences that may require further investigations. In the case of these items, this potential source of information will be lost. In addition, this elimination method may deform the content validity of the instruments.

Also, to ensure that tests are fair for all examinees, most large testing programmes have a formal review, which is part of the test development process where items are screened by content specialists for test that might be inappropriate or unfair to relevant subgroups in the test-taking population including female examinees. These reviews are conducted before the tests are administered. Statistical measures of DIF can also help test developers identify item that are biased against examinees since the ultimate criterion of item equivalence must come from an analysis of the examinee responses (Gierl, Khaliq & Boughton, 1999). According to them, the purpose of review is to examine the tests for content validity, curricular validity, item appropriateness (e.g. wording, length, interest), bias (e.g. gender, cultural, disability), balance to the test blueprint and tone.

Consequently, if characteristics that an item measures are relevant to the trait in the source but not in the target culture, such culturally specific characteristics or concepts are referred to as emic, and in contrast culturally general concepts are called as etic concepts (Hui &Traindis, 1985; Hulin & Mayer, 1986; Hulin, 1987). The general aim in cross-cultural studies is increasing the sensitivity and cultural relevance of the instrument for both cultures, but at the same time retaining the psychomotor equivalence. For this aim, etic items can be used as anchor items in linking the two language forms of a test to generate scales that only reflect cultural uniqueness but at the same time satisfying the psychomotor equivalence (Sireci, 1997).

7. The Use of Culturally Responsive Instructional Technique to Reduce Students' Differential Performance in Mathematics

Cultural psychologist and child development researchers and theoreticians have acknowledged that culture and society play a critical role in cognitive development (Vygotsky, 1979; Wertsch, Del Rio & Alvarez, 1995). Thus, culture influences the way in which people construct knowledge and create meaning from experience (how they think about things, reason and solve problems) (Greenfield, 1997), which relates directly to the ways in which individuals learn and teach informal and school settings (Lipka, 1991). The implication is that the child's culture can be used to teach him mathematics and this may make mathematics teaching and learning meaningful. According to Enukoha, the "Okoso" game or "Top" could be used to gain insight into geometry and physics therefore affording the child and teacher an opportunity for interdisciplinary discussion (Enukoha, 1995).

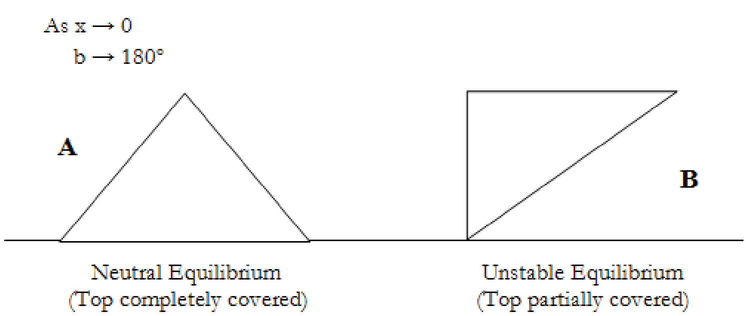
8. The Top or Okoso Game

An interesting and mathematical aspect of spinning the top (okoso) is the controversy which is often generated about the winning position of a top which is not completely covered at certain angle of inclination to the ground (horizontal). For example A and B below are two positions of a top which was covered by two players. "A" has attained neutral equilibrium while "B" has attained unstable equilibrium. The position of B generates a lot of arguments as to whether it should be accepted as a winning position or not. The basis for detecting the acceptability is the size of the angle of inclination x. The smaller the x, the more likely that the position will be accepted as a winning position (Enukoha, 1997). The children could convert these arguments into geometry class involving measurement of the various sizes of X and the relationship between x, b and w in the meaning of complementary and supplementary angles where:

$x + w = 90^o$ (Complementary angles)

$x + b = 180^o$ (Supplementary angles)

This game also give the teacher an opportunity for an interdisciplinary lesson in physics by relating the various equilibrium position (neutral and unstable) to the mathematical concepts of complementary and supplementary angles (Enukoha, 1997). For instance the bigger the x, the smaller the b and the more stable the top become until it approaches a neutral equilibrium position when x = 0.



9. The Teacher's Role

This game could be easily adapted in the class with the teacher facilitating the decision as regards acceptable winning positions. After the children have played the game on a sand tray, a number of measurements of x are made. The children and the teacher then decide what the largest x that could be accepted as a winning should be. Once this decided, the argument about winning positions is settled and the game could then become more educational having rules and standards for winning (Enukoha, 1997). According to him, there are number of these traditional games, puzzles and rhymes which could be used to teach certain mathematical concepts and process and if these are developed they may help to make mathematics teaching and learning culturally relevant.

10. Nomadic and Migrant Fishermen Activity

An illustration of the type of a mathematics problem that is culturally relevant in nomadic evaluation appeared in Pythagoras of April 1995. The title of the problem is "the heritage of fractions" and is presented below:

> *A cattle breeder owned 17 cows. His last will said: the eldest son will have ½ of them, the second 1/3 and the youngest 1/9 of them. After his death the cows had to be divided. The sons could not solve the problem without killing the cows, which according to the will was forbidden, so they started quarrelling. A wise man came along, asked what was wrong, heard the problem and asked them to wait. "I will be back in a few minutes", shortly after that, the man came back with a cow and added it to the little herd. "Well" he said, "each of you can now take your share". The eldest son took ½ of 18 = 9 cows, the second son took ⅓ of 18 = 6 cows and the youngest son took 1/9 of 18 = 2 cows. The wise man took the cow, which he borrowed from a neighbor and brought it back, leaving each of the sons satisfied.*

The problem above according to Enukoha (1997) illustrates an ingenious method of solving a problem in fraction. The problem itself is culturally very relevant to the nomad whose daily living revolve on cattle. The solution to the problem, he observed, also gives an insight into the meaning of approximation in mathematics problem solving. For instance, the problem has been solved by the use of approximations and the same result will be got but the solution would have remained at the abstract level.

½ of 17 = 8.5 = 9 cows
⅓ of 17 = 5.7 = 6 cows
$^{1}/_{9}$ of 17 = 1.9 = 2 cows

The lives of the nomad revolve around cows in the same way that the lives of fishermen revolve around fishes. Some mathematical concepts that could be developed with migrant fishermen may include: shapes, for example the conical shape of the fishing traps; patterns on the fishing nets; weight and volume of the fishing boats as well as the relationship between weight, volume, density and flotation (Enukoha, 1997). In fact, research has gingered great response of scholars to the emergence of mathematical thinking in the local parlance, using local examples of things that can be seen, recognized in their quantities as well as their similarities in shapes and cultural applications and remembered as many times as they need to be recalled. The teacher of mathematics will find it very useful as a means of motivation to let his pupils see the value of native history and past history of mathematics that can open the door to enrichment of mathematics (Odili & Okpobiri, 2011).

11. Cultural Mathematics of the Igbo Boys

Here, the boys who are already apprenticed to various trades, such topics as interest rates, simple interest, profit and loss, fractions, percentages and exchange rates of different currencies of the world, may enhance their knowledge of such topics in mathematics. The use of mathematics in the form will meet the cultural needs of these boys. It is therefore, strongly hoped that the use of these cultural materials in teaching the regular school mathematics programme will counteract with euro-centric bias recently encountered in mathematics (Enukoha, 1997). Cultural mathematics helps link past experiences to classroom experiences, links home background to the classroom activities thereby making mathematics concrete and a reality due to its processes. As such, it is pertinent that a thorough understanding of any cultural group's worldview can expose elements of teaching within their worldview which can be applied creditably in the teaching and learning of mathematics in that group (Odili & Okpobiri, 2011).

12. Conclusion

An attempt has been made in this paper to discuss the cultural imperatives of DIF. In this paper, the writer examines the concept of DIF and culture, as well as sources of DIF in multi-cultural assessment and how to deal with inequivalent items. Furthermore, the cultural responsive techniques to reduce students' differential performance in mathematics were examined.

References

Allalouf, A; Hambleton, R.K. & Sireci, S.G. (1998). Identifying the causes of DIF in translated verbal items. *Journal of educational measurement*, 36 (3), 185 – 198.

Arim, R.G. & Ercikan, K. (2005). Comparability between the US and Turkish versions of the Third International Mathematics and Science Study's Mathematics Test Results. Paper presented at NCME April 12 – 14, Montreal, Canada.

Bauersfeld, H. (1995). Language games in the mathematics classroom, their function and their effects. In P. Cobb and H. Bauersfeld (Eds) *The emergence of mathematics meaning: Interaction in classroom cultures*. Hillsdale: Lawrence Erlbaum Associates.

Brunner, J.S. (1985). The role of interaction formats in language acquisition. In J.P. Forgas (Ed) *Language and social situations*. New York: Springer Verleg.

Cobb, P. & Bauersfeld, H. (Ed) (1995). *The emergence of mathematical meaning: Interaction in classroom cultures*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Doolittle, A.E. & Cleary, T.A. (1987). Gender based differential item performance in mathematics achievement test items. *Journal of educational measurement*, 24 (2), 157 – 166.

Enukoha, O. (1995). *The psycho-cultural basis for teaching mathematics*. Owerri: Executive Publishers.

Enukoha, O. (1997). Cultural alternatives for teaching some primary mathematics concepts and processes. *West African journal of research and development*, 4 (3), 70 – 73.

Ercikan, K. (1998). Translation effects in international assessments. *International journal of educational research*, 29 (6), 543 – 553.

Flores, G.S. (2000). What is cultural validity in assessment? Retrieved November 2nd 2011, from http://www.edgateway.net/es/evap/print/dres/evap/news.htm

Fortune, J.C. (1985). *Choctaw comprehensive school study*. Philadelphia, M.S.: Choctaw Heritage Press.

Gierl, M.J. & Khaliq, S.N. (2000). Identifying sources of DIF on translated achievement test: A confirmatory analysis. Paper presented at the Annual Meeting of the National Council on Measurement in Education (NCME), New Orleans, Louisiana, USA.

Gierl, M; Khaliq, S.N. & Boughton, K. (1999). Gender differential item functioning in mathematics and science. prevalence and policy implications. Paper presented at the Symposium entitled Improving Large Scale Assessment in Education at the Annual Meeting of the Canadian Society for the Study of Education, Sherbrook, Quebec, Canada, June 7 – 11.

Greenfield, P.M. (1997). Culture as process: Empirical method of cultural psychology. In J.W. Berry, Y.H. Poortinga & J. Pandy (Eds) *Handbook of cross-cultural psychology (2nd Ed)*. Needham Height, M.A.: Allyn & Bacon.

Groulund, N.E. (1976). *Measurement and evaluation in teaching*. New York: Macmillan Publishing Co.

Guillerman, S. & Sharon, H. (2001). On the cultural validity of science assessment. *Journal of research in science teaching*, 38 (5), 553 – 573.

Hui, C.H. & Triandis, H.C. (1983). Multistrategy approach to cross-cultural research. The case of locus control. *Journal of cross-cultural psychology*, 16 (2), 131 – 152.

Hulin, C. & Mayer, L. (1986). Psychometric equivalence of translation of the job descriptive index into Hebrew. *Journal of applied psychology*, 71 (1), 83 – 94.

Hulin, C.L. (1987). A psychometric theory of evaluations of items and scale translations: Fidelity across languages. *Journal of cross-cultural psychology*, 18 (2), 115 – 142.

Klieme, E. & Baumert, J. (2001). Identifying national cultures of mathematics education: Analysis of cognitive demands and differential item functioning in TIMSS. *European journal of psychology of education*, 15 (3), 385 – 402.

Lam, T.C.M. (1995). Fairness in performance assessment. Eric clearing house on counselling and student service, Greensbore N.C. Retrieved in November, 2nd 2011 from http://www.ericfacility.net/ericdigest/ed311982.html

Likpa, J. (1991). Experimental, cognitive and anthropological perspectives in mathematics education. Retrieved on 27/08/2010 from http:www.une.edup-jlikpa/dl.html

Linn, R.L.; Levine, H. & Wardrop, H. (1981). An investigation of item bias in a test of reading comprehension. *Applied psychological measurement*, 5, 159 – 173.

Odili, G.A. & Okpobiri, N.R. (2011). Elechi Amadi's "The concubine": Ethnomathematics resource for teaching mathematics in Ikwerre primary schools. *African journal of education and technology*, Vol. 1 (2); 84 – 92.

Poortinga, Y.H. (1989). Equivalence of cross-cultural data: An overview of basic issues. *International journal of psychology*, 24, 737 – 756.

Radziszewska, B. & Rogoff, B. (1991). Children's guided participation in planning imaginary errands with skilled adults or peer patterns. *Developmental psychology*, 27, 381 – 389.

Roznowski, M. & Reith, J. (1999). Examining the measurement quality of tests containing differentially functioning items: Do biased items result in poor management? *Educational and psychological measurement*, 59 (2), 248 – 269.

Scheuneman, J.D. & Grima, A. (1997). Characteristics of quantitative word items associated with differential performance for female and black examinees. *Applied measurement in education*, 10 (4), 299 – 319.

Schleimann, A.D.; Carraher, D.W. & Ceci, S.J. (1997). Everyday cognition. In J.W. Berry, P.R. Dasen & T.S. Saraswathi (Eds) *Handbook of cross-cultural psychology*, 2; 177 – 216.

Sireci, S.G. & Berberoglu, G. (2000). Using bilingual respondents to evaluate translated – adapted items. *Applied measurement in education*, 13 (3), 229 – 248.

Sireci, S.G. (1997). *Problems and issues in linking assessment across languages. Educational measurement: Issues and practice*, 16 (1), 12 – 19.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, M.A.: Harvard University Press.

Warm, T.A. (1978). A primer of item response theory. Springfield, V.A.: National Technical Information Services, US Department of Commerce.

Wertsch, J.V.; Del Rio, P. & Alvarez, A. (Eds) (1995). *Sociocultural studies of mind*. New York: Cambridge University Press.

Zumbo, B.R. (1999). A handbook on the theory and methods of differential item functioning: Logistic regression modeling as a unitary framework for binary and likert type (ordinal) item scores. Ottawa, O.N.: Directorate of Human Resources Research and Evaluation, Department of National Defence. Retrieved 7/4/11 from http://www.edu.ubc.ca/faculty/zumbo/DIF/handbook.pdf